BIOTECHNOLOGY
SYSTEMS.
BRANCH

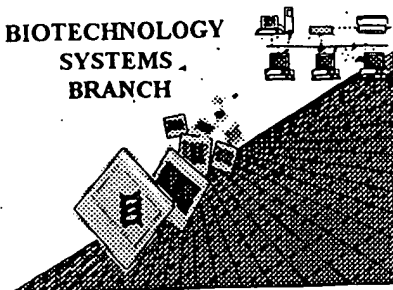# RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: *09 700 158*

Source: *OIPE*

Date Processed by STIC: *08/15/2001*

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

ERROR DETECTED   SUGGESTED CORRECTION   SERIAL NUMBER: _09/700 758_

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFT\

1_____Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2_____Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3_____Misaligned Amino Numbering

The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4_____Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5_____Variable Length

Sequence(s)____contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6_____PatentIn 2.0 "bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7_____Skipped Sequences (OLD RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8_____Skipped Sequences (NEW RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9_____Use of n's or Xaa's (NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10____Invalid <213> Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11_____Use of <220>

Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998; Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

12_____PatentIn 2.0 "bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13_____Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001

The types of errors shown exist _throughout_ the Sequence Listing. Please check subsequent sequences for similar errors.

RAW SEQUENCE LISTING   DATE: 08/15/2001
PATENT APPLICATION: **US/09/700,158**   TIME: 07:36:46

Input Set : **A:\Seq.txt**
Output Set: **N:\CRF3\08132001\I700158.raw**

Does Not Comply
Corrected Diskette Needed

see pages 1 and 2

```
 3 <110> APPLICANT: BOWNESS, RACHEL
 4        MCMICHAEL, PAUL
 5        JAMES, ANDREW
 7 <120> TITLE OF INVENTION: NOVEL MOLECULE AND DIAGNOSTIC METHOD
 9 <130> FILE REFERENCE: P02077US0/10025595
11 <140> CURRENT APPLICATION NUMBER: 09/700,158
12 <141> CURRENT FILING DATE: 2000-11-10
14 <150> PRIOR APPLICATION NUMBER: PCT/GB9901481
15 <151> PRIOR FILING DATE: 1999-05-11
17 <160> NUMBER OF SEQ ID NOS: 9
19 <170> SOFTWARE: PatentIn version 3.0
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 9
23 <212> TYPE: PRT
24 <213> ORGANISM: INFLUENZA
26 <400> SEQUENCE: 1
28 Ser Arg Tyr Trp Ala Ile Arg Thr Arg
29 1               5
31 <210> SEQ ID NO: 2
32 <211> LENGTH: 10
33 <212> TYPE: PRT
34 <213> ORGANISM: HIV
36 <400> SEQUENCE: 2
38 Lys Arg Trp Ile Ile Met Gly Leu Asn Lys
39 1               5                   10
41 <210> SEQ ID NO: 3
42 <211> LENGTH: 17
43 <212> TYPE: PRT
44 <213> ORGANISM: ARTIFICIAL PEPTIDE
46 <220> FEATURE:
47 <221> NAME/KEY: misc_feature
48 <222> LOCATION: (1)..(17)
49 <223> OTHER INFORMATION: Biotinylation Sequence
52 <400> SEQUENCE: 3
54 Gly Ser Leu His His Ile Leu Asp Ala Gln Lys Met Val Trp Asn His
55 1               5                   10                  15
57 Arg
59 <210> SEQ ID NO: 4
60 <211> LENGTH: 9
61 <212> TYPE: PRT
62 <213> ORGANISM: CMV (HLA-A2 CMV)
64 <400> SEQUENCE: 4
66 Asn Leu Val Pro Met Val Ala Thr Val
67 1               5
69 <210> SEQ ID NO: 5
70 <211> LENGTH: 26
71 <212> TYPE: DNA
```

C--> (line 44)

Errored. The appropriate 213 response is Artificial Sequence ✓

The types of errors shown exist <u>throughout</u> the Sequence Listing. Please check subsequent sequences for similar errors.

```
C--> 72 <213> ORGANISM: ARTIFICIAL    ✓
     75 <220> FEATURE:
     76 <221> NAME/KEY: misc_feature
     77 <222> LOCATION: (1)..(26)
     78 <223> OTHER INFORMATION: DNA primer  ✓
     81 <400> SEQUENCE: 5
     82 gagacacaga tcagcaaggc caaggc                                    26
     85 <210> SEQ ID NO: 6
     86 <211> LENGTH: 26
     87 <212> TYPE: DNA
C--> 88 <213> ORGANISM: ARTIFICIAL    ↺
     91 <220> FEATURE:
     92 <221> NAME/KEY: misc_feature
     93 <222> LOCATION: (1)..(26)
     94 <223> OTHER INFORMATION: DNA Primer  ✓
     97 <400> SEQUENCE: 6
     98 gccttggcct tgctgatctg tgtctc                                    26
    101 <210> SEQ ID NO: 7
    102 <211> LENGTH: 21
    103 <212> TYPE: (DNA)
C--> 104 <213> ORGANISM: ARTIFICIAL
    107 <220> FEATURE:
    108 <221> NAME/KEY: misc_feature
    109 <222> LOCATION: (1)..(21)
    110 <223> OTHER INFORMATION: ARTIFICIAL (PEPTIDE)
    113 <400> SEQUENCE: 7
    114 tttgttgaat tcaggaggaa t
    117 <210> SEQ ID NO: 8
    118 <211> LENGTH: 67
    119 <212> TYPE: DNA
C--> 120 <213> ORGANISM: ARTIFICIAL    ✓
    123 <220> FEATURE:
    124 <221> NAME/KEY: misc_feature
    125 <222> LOCATION: (1)..(67)
    126 <223> OTHER INFORMATION: DNA Primer
    129 <400> SEQUENCE: 8
    130 cggggagtgg gactctaccc tccctaggga cgtagtataa gacctacgtg tcttttacca   60
    132 cacctta                                                           67
    135 <210> SEQ ID NO: 9
    136 <211> LENGTH: 22
    137 <212> TYPE: PRT
C--> 138 <213> ORGANISM: ARTIFICIAL    ✓
    141 <220> FEATURE:
    142 <221> NAME/KEY: misc_feature
    143 <222> LOCATION: (1)..(22)
    144 <223> OTHER INFORMATION: Fusion protein  ✓
    147 <400> SEQUENCE: 9
    149 Pro Leu Thr Leu Arg Trp Glu Gly Ser Leu His His Ile Leu Asp Ala
    150 1               5                   10                  15
```

Handwritten annotations:

✓ DNA is by definition a nucleotide sequence

you have entered artificial peptide sequence and have failed to describe your artificial nucleotide sequence in field 223.

RAW SEQUENCE LISTING
PATENT APPLICATION: **US/09/700,158**

DATE: 08/15/2001
TIME: 07:36:46

Input Set : **A:\Seq.txt**
Output Set: **N:\CRF3\08132001\I700158.raw**

152 Gln Lys Met Val Trp Asn
153   20

VERIFICATION SUMMARY   DATE: 08/15/2001
PATENT APPLICATION: **US/09/700,158**   TIME: 07:36:47

Input Set : **A:\Seq.txt**
Output Set: **N:\CRF3\08132001\I700158.raw**

L:44 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:3
L:72 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:5
L:88 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:6
L:104 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:7
L:120 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:8
L:138 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:9